R E P O R T R E S U M E S

ED 017 466 SE 004 229

A FEASIBILITY STUDY FOR DETERMINING REQUIREMENTS OF BIOLOGICAL INFORMATION SERVICES AND SYSTEMS.

BY- VAN COTT, HAROLD P. KINKADE, ROBERT G.
NATIONAL SCIENCE FOUNDATION, WASHINGTON, D.C.

REPORT NUMBER NSF-C492 PUB DATE NOV 67
EDRS PRICE MF-$0.25 HC-$0.88 20P.

DESCRIPTORS- *INFORMATION CENTERS, *INFORMATION NEEDS, *INFORMATION SYSTEMS, BIOLOGY, COLLEGE SCIENCE, INFORMATION STORAGE, SCIENTISTS, SECONDARY SCHOOL SCIENCE, SCIENCE INFORMATION SERVICE, AMERICAN INSTITUTES FOR RESEARCH,

EVALUATED WAS THE FEASIBILITY OF USING A PILOT INFORMATION CLEARINGHOUSE AS A RESEARCH TOOL FOR DEFINING DESIGN REQUIREMENTS FOR AN ADVANCED BIOLOGICAL INFORMATION CENTER. SPECIFIC OBJECTIVES WERE (1) TO DEVELOP A METHOD FOR STUDYING THE SCIENCE INFORMATION REQUIREMENTS OF BIOLOGICAL SCIENTISTS, (2) TO EVALUATE THE FEASIBILITY OF USING THE METHOD AS A RESEARCH TOOL TO DETECT CHANGES IN THE BEHAVIOR OF USER SCIENTISTS AS A FUNCTION OF MODIFICATION IN CLEARINGHOUSE PERFORMANCE, AND (3) TO GENERATE TENTATIVE PRESCRIPTIONS FOR SYSTEMS DESIGN. THE PILOT INFORMATION CLEARINGHOUSE WAS LOCATED AT THE FEDERATION OF AMERICAN SOCIETIES FOR EXPERIMENTAL BIOLOGY. THE RESULTS SHOWED (1) THAT MEANINGFUL DATA COULD BE OBTAINED CONCERNING THE REQUESTING BEHAVIOR OF SCIENTISTS, AND (2) THAT THERE WERE CHANGES IN THE BEHAVIOR OF SCIENTISTS AS A RESULT OF CERTAIN ELEMENTS IN THE CLEARINGHOUSE PROGRAM. TWO SUPPORTING STUDIES WERE CONDUCTED. THE FIRST WAS DESIGNED TO DETERMINE THE POSSIBILITY OF INCREASING THE EFFECTIVENESS OF THE PILOT INFORMATION CLEARINGHOUSE BY USING EXISTING INFORMATION SYSTEMS. THE SECOND STUDY WAS DESIGNED TO PROVIDE INFORMATION ABOUT WHAT ADDITIONAL SERVICES BIOLOGICAL SCIENTISTS MIGHT FIND USEFUL. THE RESULTS OF ALL THE STUDIES INDICATED (1) THAT IT WOULD NOT BE PARTICULARLY EFFECTIVE TO USE EXISTING INFORMATION SYSTEMS TO FACILITATE THE OPERATION OF THE PILOT INFORMATION CLEARINGHOUSE, AND (2) THAT A DESIRABLE SERVICE WOULD BE TO PROVIDE SCIENTISTS WITH "CURRENT AWARENESS" INFORMATION. (DS)

.. 41229

234

AIR-F-57-11/67-TR-7

ED017466

SCIENCE INFORMATION REQUIREMENTS OF SCIENTISTS--VII:

A Feasibility Study for Determining Requirements of Biological Information Services and Systems

Harold P. Van Cott

Robert G. Kinkade

FINAL REPORT

Prepared under Contract for
Office of Science Information Service
National Science Foundation

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Contract No. NSF-C492

American Institutes for Research
Washington Office
Skills Research Program

In Cooperation With

The Federation of American Societies for Experimental Biology

November 1967

SE004 229

FOREWORD

The investigation described in this report was supported by the National Science Foundation, Contract Number C 492.

As part of the project described in this report, an Information Clearinghouse was established at the Federation of American Societies for Experimental Biology. Dr. Beverly L. Clarke, Research Officer at the Federation, served as a Co-Principal Investigator on the project, and was responsible for organizing and operating the Clearinghouse. Working with Dr. Clarke were Dr. Henry Stevens, recently retired as Chief Biochemist, Allergins Laboratory, U. S. Agricultural Research Service, and Mrs. Grace Gurtowski, previously with the Federation's Translation Project. Also assisting on a part-time basis were Mrs. Ruth Katz, Science Division of the Library of Congress, and Mrs. Ruth Stander, formerly with the National Cancer Institute.

M. G. Fromm, G. R. Wheaton, Catherine B. Judd, and W. P. Gloege contributed to various phases of this study.

In addition, the cooperation of the biological scientists who participated in this study is gratefully acknowledged.

I. OVERVIEW AND SUMMARY

The American Institutes for Research evaluated the feasibility of using a pilot information clearinghouse as a research tool for defining design requirements for an advanced biological information system. The pilot information clearinghouse was operated at the Federation of American Societies for Experimental Biology (FASEB). At the clearinghouse an experienced team of biologists, psychologists and library scientists was available to receive, process, fill and analyze requests for scientific information as phoned in by biological scientists working in the Washington, D. C. area.

The clearinghouse resembles the properties that could be designed into an improved manual or automatic information systems by using humans to represent alternative manual or automatic functions. The use of a pilot information system as a research tool provides the means for collecting empirical data regarding the information-related activities of scientists and the means for identifying and evaluating system design alternatives.

Three studies involving the operation of the pilot information clearinghouse were conducted. The first study was designed to answer questions concerning the kinds of services or functions that scientists would demand of a clearinghouse and the way in which these demands would be made. The second and third studies were designed to answer questions concerning the ability of the research tool to detect changes in the behavior of the scientists as a function of modifications in the performance of certain elements in the clearinghouse.

In addition to these three studies, two supporting studies were conducted. The first of these studies was designed to determine the possibility of increasing the effectiveness of the pilot information clearinghouse by using existing information systems. The second study was designed to provide information about what additional services biological researchers might find useful.

The results of the studies concerned with evaluating the use of the pilot information clearinghouse as a research tool showed that meaningful data could be obtained concerning the requesting behavior of scientists and that changes in the behavior of scientists as a function of modifications in the performance of certain elements in the clearinghouse could be detected. The results of the supporting studies showed that it would not be particularly effective to use existing information systems to facilitate operation of the pilot information clearinghouse and that a desirable service, from the viewpoint of the biological researcher, would be to provide the scientists with "current awareness" information.

II. BACKGROUND

There is a growing interest in the design and implementation of automatic information systems. It is felt that such information systems will increase the quality of research that is being conducted, replace time that is currently spent for information search with time for laboratory research, and permit scientists to gain better access to the ever increasing volume of scientific literature. Before systems which will accomplish these functions can be designed, it is necessary to determine which of many possible alternative design features should in fact be implemented (Salton and Lesk, 1967). Guidelines for specifying design alternatives are now lacking because very little is known about the information requirements of scientists (Paisley, 1965). The purposes of the research described in this report were to: (1) develop a method for studying the science information requirements of biological scientists; (2) evaluate the feasibility of using the method as a research tool; (3) provide hypotheses about the information related behavior of biological scientists; and (4) generate tentative prescriptions for systems design.

The choice of a method for identifying the requirements of scientists for an information system is a critical one. A recent review of studies concerning information needs and uses in science and technology (Herner and Herner, 1967) failed to cite even one study using an experimental method. The methods that were used included diary, interview, passive observation, questionnaire, bibliographic analysis, or combinations of these methods. Although these approaches contribute some insights into the information requirements of scientists, they either intrude on the behavior of the scientist to change it or may confuse what a scientist says he needs or wants with what he actually uses. The information seeking and using behavior of the scientist must be studied more objectively and directly in the context of daily research activities in order to obtain valid design information.

This could be done by establishing a pilot information system and observing the scientists' behavior while using the system. Such a system would have to possess flexibility in order to adapt to its users rather than forcing the users to adapt to it. By assembling human elements into an information system, where each element has an assigned function, it should be possible to create a large number of alternative models by varying the number of human elements, their communication links, training and instructions. The behavior of each configuration can then be measured experimentally, comparing it with other models or with performance criteria established by its users. But most importantly, by placing few or no constraints on its users, those features that the system acquires as it adapts to its users could be identified and observed.

One feature of this approach needs to be made explicit. It would provide the means for bridging the gap between so-called user studies and system design requirements studies. The compartmentalization of these two types

of studies is exemplified in the 1967 Annual Review of Information Science and Technology in which discussions of "user requirement", system analysis, and system evaluation are handled as independent topics; each having its own distinctive literature (Cuadra, 1967). At no place is it suggested that "users" can be studied by way of experimental methods, that analysis depends on data obtained by way of systematic observation; or that the ultimate criterion for the evaluation of systems is not "relevance assessment" but some direct measure of user satisfaction. The approach evaluated in this project provides a direct means for empirically evaluating user and system design requirements simultaneously.

Two different types of studies are reported. One concerns the feasibility of using a pilot information clearinghouse as a research tool. The other concerns supporting research related to establishing the conditions for continuing operation of the pilot information clearinghouse. Each type is reported in separate sections.

III. FEASIBILITY EVALUATION STUDIES

A. Objectives

The primary goal of the present research was to evaluate the feasibility of using a pilot information clearinghouse as a method for studying the information requirements of biological scientists. In this context, the clearinghouse was analogous to the apparatus in a laboratory experiment.

In pursuit of this goal, three basic sets of questions had to be answered. First, could such a facility be brought into being with the resources available? Second, in addition to serving as a research instrument, would such a facility actually operate in the sense of providing meaningful services to subject scientists? The third and final question concerned whether it would be possible to control conditions and manipulate important aspects of the service in such a way as to be able to draw conclusions about the impact of change and innovation on the behavior of scientists. In other words, would such a pilot information clearinghouse serve as a sufficiently sensitive research instrument?

To answer these and similar questions, the plan was to establish the information clearinghouse and allow it to operate in one basic mode for a period of time in order to produce a reasonably stable pattern of transactions. Following the establishment of such a "base-line" condition and the exploration of various observational procedures and dependent measures, a series of manipulations of operating conditions and procedures could be undertaken as modest "experiments."

The central theme of the project, therefore, consisted of a step toward the experimental determination of design requirements for systems providing information services to scientists. The pilot system represented the core research instrument for carrying out the theme.

B. Description of the Pilot-Information Clearinghouse

The overall structure of the facility was derived from a synthesis of the traditional ideas of library-reference service with the more recent concepts of information processing and dynamic simulation. The ultimate product of such a synthesis might well be a computerized reference and retrieval system or it could be an improved manual system. The putative gains from either type of system would be rapid response combined with ease of access and individualized service (Rubinoff, 1965; NINDB, 1966; Schecter, 1967).

One way of exposing the processes in an advanced information system to empirical observation would be to block out critical sub-functions and

assign these sub-functions to human operators. This is what was done in the establishment of the pilot information clearinghouse. The functional structure is summarized in Figure 1.

This functional structure was implemented at the Federation of American Societies for Experimental Biology (FASEB). Trained, biological scientists with many years of experience in research were available to serve as functional elements in the system. Functional elements requiring a knowledge of the library sciences were represented by individuals possessing extensive experience in performing library services. The operation of the clearinghouse was administered by a member of the FASEB staff who could draw upon the resources of the organization for consultation when needed.

Biological scientists in the Washington, D. C. area were invited to use the clearinghouse to obtain whatever scientific information they desired. By not constraining the form or content of the scientists' requests, the clearinghouse would assume the functions inherent in patterns of requests made by the subject scientists.

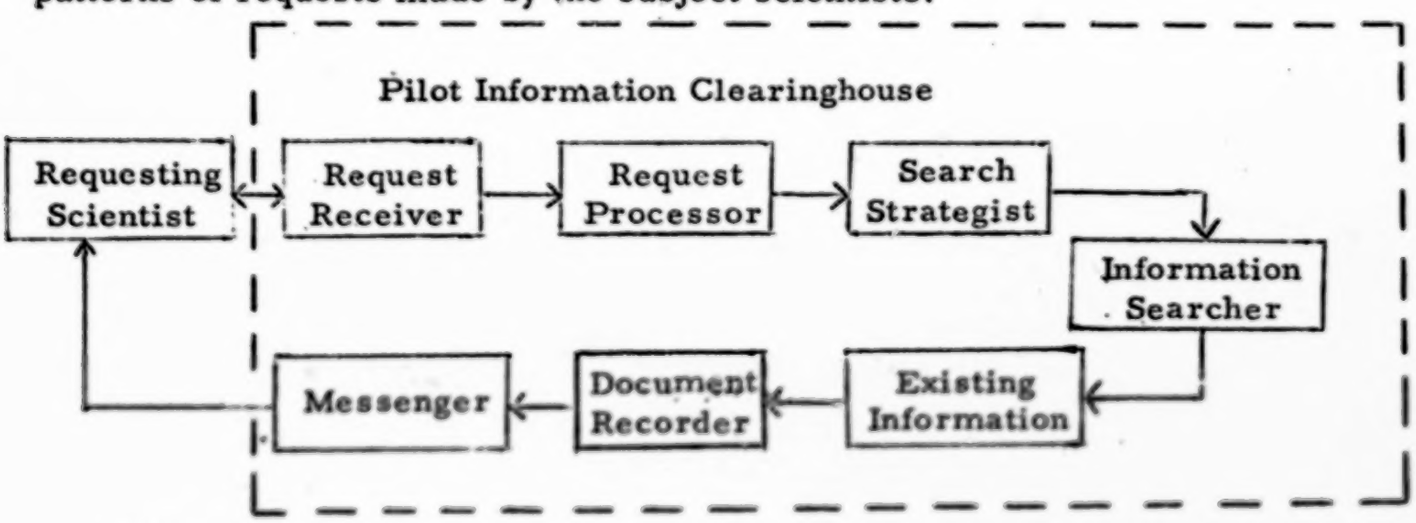

Figure 1. Representation of the Pilot Information Clearinghouse

Participating scientists made requests by calling the clearinghouse on the telephone. The call was received at the clearinghouse by the Request Receiver. The Request Receiver's task was to assure that an understandable, processable request was made by the scientists. All of the requests were recorded on a tape-recorder.

After a call had been completed, the requests made during that call were processed by the Request Processor. The Request Processor's task was to listen to the taped recording of the call and to summarize each request made during the call. In some cases, the requests were interpreted by the Request Processor who placed additional descriptors or necessary constraints on the requests so that the requested information could be obtained more effectively.

The summarized requests, as interpreted by the Request Processor, were then given to the Search Strategist. The Search Strategist's task was to decide how the requests were to be filled and to supply any additional descriptors or constraints not included in the original requests. In deciding how the requests were to be filled, the Search Strategist had to select one of several available sources.

When requests involved searching for information, the summarized requests were given to an Information Searcher, along with any additional descriptors supplied by the Search Strategist. The Information Searcher's task was to locate required information by using existing library facilities.

When the information needed to fill a request was obtained by the clearinghouse, it was logged in and then delivered to the requesting scientist by messenger. Subsequently, the scientists were contacted for their evaluations of the products supplied by the clearinghouse.

C. Evaluation Procedure

Three feasibility evaluation studies were conducted. The first study was designed to answer some major questions. (1) What kinds of services would the scientists demand of the clearinghouse? (2) Could the clearinghouse provide these services? (3) Would a reasonable number of transactions occur within the limited time frame of six weeks? To answer these questions, fifteen biological scientists were invited to use the clearinghouse. As requests were obtained, the roles of each of the functional elements in the clearinghouse were brought into sharper focus as each person (function) adapted to the requirements of the scientists.

The principal purpose of the second and third studies was to answer questions concerning the sensitivity of the scientists' behavior to changes in the functions performed by the elements of the clearinghouse. Manipulations of these functions were designed to provide insights into actual design questions that need to be answered. Design questions concerning the interface between the requesting scientists and the information system have already been raised (Swanson, 1966; Overhage, 1966). Therefore, both studies were directed toward providing data regarding the design of this interface element.

It was felt that a start toward answering such questions could be made by manipulating the characteristics of the people operating in the Request Receiver and Request Processor positions. Therefore, in the second study a receptionist-stenographer replaced the trained biologist as both Request Receiver and Request Processor on a systematic schedule. The subject-scientists were informed of the Request Receiver schedule so they knew in advance whether a telephoned request for service would put

them in contact with an experienced biologist or a clerically-trained lay person. They were not told, however, that the lay person would process half of their requests because this might bias their evaluations of the products supplied by the clearinghouse.

In the third study of the pilot information clearinghouse, the focus of manipulation continued to be the Request Receiver component but the differentiation between alternative modes was made more pronounced. The question at issue became the nature of the conversational feedback presented to the scientist making a request. In some prospective versions of highly mechanized systems, conversational feedback might be totally absent or artificial. The positive incentive in such systems would be almost instant access (e.g., COSATI, 1965), but some scientists might have an aversion for such an impersonal transaction (Taylor, 1962). Two Request Receiver types were evaluated. One Request Receiver type was represented by a trained, experienced biological scientist and the other was represented by a tape-recorder. The recorder-receiver gave no feedback to the requesting scientists, but only recorded their requests. The scientist-receiver provided both "technical" and "conversational" feedback. The requesting scientists were told that their requests would be received by these two types of receivers and they were given a schedule that showed when each type would be receiving calls.

In the second study, forty-six scientists were invited to participate, including the original fifteen in the first study. In the third study, an additional five scientists were invited to participate. Observations were made during a six week period in the second study. Vacation schedules interfered with the third study and observations were made only during a five-week period.

D. Results

In the first study, all fifteen scientists who were invited to use the pilot information clearinghouse placed at least one call. Within the six-week period 101 requests were made during 34 calls to the clearinghouse. A request was defined as any statement made by the subject-scientists to which the clearinghouse could respond with a document or a written report. Although full participation was not achieved during the other two studies, where substantially larger subject samples were employed, the number of requests received by the clearinghouse remained at an acceptably high level. Over the entire 17 week period of clearinghouse operation, a total of 372 separate requests were placed, permitting examination of requesting behavior for a variety of dependent variables.

In each of the clearinghouse studies there was pronounced variability in the requesting behavior of the scientists. For example, the number of requests made during a single call was highly variable, ranging from one

to 13 requests per call. Similarly, no consistent pattern emerged regarding the timing of calls. Peak loads occurred during certain hours, days, and weeks but these peak periods were not necessarily consistent from one study to another. Individual differences in the information needs of scientists were shown by the decline in active participation by some scientists and the continued participation by others. For example, thirty scientists actively participated in only one of the three studies, and these scientists made, on the average, only 3.7 requests during that study. Fifteen of the scientists actively participated in two studies and they made an average of 4.7 requests per study. Only three of the original fifteen scientists actively participated in all three studies. These three scientists made an average of 14.3 requests per study. Taking all the scientists who were invited to use the clearinghouse into account, fewer than 5 per cent of the scientists accounted for over 25 per cent of the requests received. The absolute values of these figures are confounded by the fact that only the original fifteen scientists had an opportunity to participate in all three studies and that there were time lapses between studies when some scientists may have used other sources to fill their information needs. However, the trends are clear - some scientists will be continuing heavy users of an information system and some will use the system infrequently.

In general, the subject-scientists requested two types of information: specific documents and bibliographic searches. Combining the results of all three studies, the ratio of document requests to bibliographic search requests was approximately 2:1. The principal focus of bibliographic search requests was experimental findings regarding a specific subject. About two-thirds of the bibliographic search requests were posed in terms of such findings. The remainder were predominantly concerned with methodological content. Very few requests from the subject-scientists concerned theoretical information.

When requesting documents, the scientists were very specific in the descriptors they used, permitting easy retrieval of the document. Titles, authors, dates, etc. were accurately supplied in almost all cases. When requesting bibliographic searches, the scientists supplied a number of key-word descriptors. However, performing a bibliographic search using these descriptors was far from routine.

In the first study it took an average of seventeen days to fill a request. However, the time to fill a request varied from one to thirty-eight days. The longer delays were non-essential, and may be regarded an an artifact. They were caused by initial personnel problems and an initial backlog in filling requests rather than the response time of information sources. In general, it took at most two weeks to obtain a copy or portion of a document from either a facility that provides Xerox copies or from publishers. Neglecting non-essential delays it took about a week to fill requests for information involving a bibliographic search.

The scientists' attitudes toward the pilot clearinghouse and its operation, as expressed during interviews following the first study, are summarized in Table 1. The table shows an assessment of the value associated with scientists' comments concerning several topics. Three

Table 1

Assessment of the value associated with comments made by scientists concerning the system and its operation

Value Associated with Comments

Topic	Very Negative	Negative	Neutral	Positive	Very Positive	Total
Concept of the system	0	0	2	6	7	15
Use of Scientist to receive calls	0	0	0	5	10	15
Usefulness of info. received	0	2	1	6	6	15
Completeness of info. received	1	3	2	6	3	15
Timeliness of Information	1	0	3	10	1	15
Time required to fill request	0	9	2	4	0	15
TOTAL	2	14	10	37	27	90

items are of particular importance. First, the general reactions to the pilot operation were favorable both in terms of the products supplied and in terms of the clearinghouse concept. The second important reaction, and one which suggested the experimental plan for the second and third studies, concerned the interfacing function between the system and its users. Scientists were asked what they thought of having a trained biological researcher serve as a Request Receiver. All of the scientists responded positively to this condition. In fact, some of the participants felt that it was absolutely essential. The third item of particular importance concerned the scientists' feeling about the time required to fill the request. In general, the scientists felt that the delay in receiving a product from the clearinghouse in the first study was too long. In the subsequent studies, the time required to fill a request was reduced to an average of seven days because of improved clearinghouse administrative procedures, and this was found to be acceptable to nearly all of the participating scientists.

Summarizing the results of the first study, it was found that scientists would use a pilot information clearinghouse and that the clearinghouse could provide meaningful services in response to the requests made by scientists. A sufficient number of transactions occurred during operating periods that relatively stable descriptive data concerning the requesting behavior of scientists could be obtained.

Turning now to the data concerning the ability to detect changes in the behavior of the scientists as a function of experimental manipulations, it will be recalled that the primary focus of these experimental manipulations was the interface between the subject-scientists and the clearinghouse. Changes in the behavior of the scientists were detected as a function of these experimental manipulations.

An interaction between the type of receiver, i.e., trained biological scientists vs. an inexperienced receptionist, and what the scientists ask for was found. In terms of conditional probabilities, if a subject-scientist wants a document, there is roughly an equal probability of choosing either receiver type (0.52 for the scientist and 0.48 for the receptionist). However, if the subject-scientist wants a bibliographic search performed, there is a greater probability (0.74) that the scientist will be called.

There did not appear to be any significant differences between the scientist and receptionist in terms of the time required to process a request, or the subject-scientists' evaluation of the product supplied by the clearinghouse. However, the receptionist was unable to process four of the calls assigned to her. The scientist easily processed these calls.

In the subsequent (third) study, differences between receiver type, i.e., biological scientist vs. tape-recorder, and type of request were again

detected. In terms of conditional probabilities, if a subject-scientist wants a bibliographic search conducted, he is more likely to place the call with the scientist-receiver (p=0.64) than with the recorder-receiver (p=0.36).

However, an additional difference was obtained. If a subject-scientist wants a document, the probability of calling the scientist-receiver is higher (p=0.66) than the probability of calling the recorder-receiver (p=0.34). This difference between receiver types was not obtained in the second study when documents were requested.

In interviews following each of the studies, most of the subject scientists stated that they attended to the request receiver schedule mailed to them each week. They consciously contacted the scientist-receiver when they felt that their requests could not be easily articulated. In addition, they felt "uncomfortable" when they placed a call with the recorder-receiver, even when they had a well articulated request, because they were never sure that their requests were "understood".

D. Discussion

The principal goal of this research project was to demonstrate the feasibility of using a pilot information clearinghouse as an experimental instrument for studying the information requirements of biological scientists. To accomplish this goal, an inter-disciplinary team composed of biological scientists, library scientists and psychologists combined their talents in the establishment, operation, and evaluation of the pilot information clearinghouse.

The most important finding of this research was that members of the biological-scientific community will use the service offered by a pilot information clearinghouse. A sufficient number of transactions occurred over a six-week period of time, using a relatively small subject population, so that consistent and reliable results concerning scientists' information requesting behavior were obtained. In addition, changes in the subject scientists' behavior were observed when relatively minor adjustments in the functions performed by elements in the clearinghouse were made. The method is sensitive to different experimental treatments.

Based on the data obtained during these studies, certain hypotheses concerning scientists' requesting behavior that have implications for system design, can be made. It may be hypothesized, for example, that there will be brief periods of peak activity and long periods of relative idleness in an information system serving biological researchers. Either the system must be designed to handle these peak loads and remain idle the majority of the time (or perform other tasks) or constraints must be

placed on the scientists which will cause them to space their requests more evenly. From a cost-effectiveness standpoint, the operant level of requesting behavior should be explored in greater detail, including the load that could be expected during difficult months and seasons of the year.

Another hypothesis based on the observed variability in the requesting behavior of the scientists sampled can be made. It is predicted that at least some scientists will tend to "save up" their requests and output all of these requests during a single contact with the information service. The system would have to be designed so that it could recognize when the requesting scientist had finished making one request and started making another. This might be particularly difficult for an automated system which handles requests for bibliographic searches. Again, however, constraints on the scientists requesting behavior could obviate the requirements for such a recognition capability, providing they did not unduly constrain the scientist user.

There were certain aspects of the scientists' behavior which were consistent. Scientists were invited to use the clearinghouse to obtain any scientific information that they desired. In general, they restricted their requests to two types: specific documents and bibliographic searches. In almost all cases, the scientists provided complete and accurate descriptors when asking for specific documents. Therefore, it appears that searches for documents which are only partially referenced are not very likely to occur and it may not be necessary to design a system to perform such searches. It is <u>not</u> hypothesized that a system providing only documents or bibliographic searches will be totally satisfactory. It appears that the scientists arbitrarily placed these constraints on their requesting behavior because they thought that the pilot operation could perform only these two types of services.

Most scientists judged a delay in receiving the requested products from the clearinghouse on the order of two to three weeks as being unacceptable. However, a delay on the order of one to two weeks was judged to be completely acceptable by almost all of the scientists. Therefore, it is hypothesized that it may not be necessary to design a system with an immediate display capability to obtain user satisfaction.

The data resulting from the studies which assessed the impact of minor modifications in the functions performed by the Request Receiver element show that some interactive capability on the part of the interface between the requesting scientist and the information system is necessary. It appears that this interactive capability should involve two types of feedback: technical and conversational. Technical feedback is required to help the requesting scientist formulate his request. Conversational feedback is required to indicate that the request is understood

or, at least, has been accepted. As a tentative prescription for system design, therefore, the interface element should be capable of providing these two types of feedback.

In conclusion, the feasibility of using a pilot information clearinghouse as a research tool has been demonstrated. The primary goal of this research was to demonstrate the feasibility of using such a clearinghouse as a research tool and this goal has been achieved. In addition, certain hypotheses regarding the behavior of scientists which have design implications can be extracted from the data obtained. Also, tentative design requirements concerning the nature of the interface between the scientist and the information system can be stated. Further research, using the pilot information clearinghouse with larger samples of subject-scientists, is needed to verify these hypotheses and prescriptions.

Finally, it appears to be possible to use humans to represent functions which could be performed by automatic components in operational systems -- a method which may be demonstrated to be more flexible and less costly as an R and D approach than using equipment to represent these functions.

IV. SUPPORTING STUDIES

A. Objectives

It was explicitly stated at the outset of this project that it was the first step in a continuing program to establish the design requirements of an advanced information system serving biological scientists. Therefore, once the feasibility of using the pilot information clearinghouse as a research tool had been established, two supporting studies were conducted. The overall objective of these studies was to supply information that was needed for future research using the pilot information clearinghouse.

The purpose of the first study was to define the characteristics of existing information services related to biology in the Washington, D. C. area. It was felt that it might be possible to take advantage of these existing services in the operation of the pilot clearinghouse, thereby providing more effective and varied services to the requesting scientists.

The purpose of the second study was to determine the attitudes of scientists who are performing basic research in the field of biology toward information systems. Specifically, the second study attempted to answer the following kinds of questions; do these scientists feel that they take full advantage of the services offered by existing information systems? Do the systems adequately meet their perceived informational needs, or are additional services required? What other services do the scientists believe they would use?

B. Procedure

Scientific information systems operating in Virginia, Maryland, Delaware, and the District of Columbia were reviewed. Twelve systems were ultimately selected for inclusion in the sample to be studied further. After selection, a senior-staff representative from each facility was interviewed by telephone and, in addition, descriptive brochures were obtained from nine of the twelve facilities. The descriptive information concerning these systems was then classified into eleven "characteristic" categories.

Information about biological researchers' attitudes toward existing information systems was obtained using partially structured telephone interviews. The intent of these interviews was to obtain qualitative, descriptive information concerning the scientists' attitudes.

Fifty biological researchers meeting specified criteria participated in the study. To be invited to participate, the scientists had to be (a) actively engaged in conducting or directing basic biological research, (b) recognized as a competent researcher by other biologists, and (c) employed on a full-time basis by an academic institution, i. e., college or university. In addition to these restrictions, an attempt was made to select a

representative sample of biological researchers with respect to (a) size of the institution employing the scientists, (b) distance of the institution from Washington, D. C., and (c) the scientists' area of specialization within biology.

C. Results and Discussion

1. Survey of existing information systems.

Although the survey of existing information systems included the specification of eleven characteristics for each system, only two are of particular concern to the immediate goals of the project: patron qualifications and access procedures. Only two services are open to the general public, neither of which conducts general bibliographic searches. The Clearinghouse (CFSTI) has discontinued this service and the National Referral Center only provides clues to the location of information in the form of names of people and institutions that can be associated with a particular research topic or problem.

Only nominal credentials are required, however, for access to four of the sampled facilities. Two of these provide bibliographic search and retrieval services, but the subject area is narrowly restricted (Automatic Data Processing Service of the Armed Forces Institute of Pathology and the Cancer Chemotherapy National Service Center). A third provides notification of work-in-progress and potential sources, but does not yield lists of documents (Science Information Exchange). Moreover, the SIE covers only government sponsored work and it is widely acknowledged that this coverage is not exhaustive. The National Library of Medicine--MEDLARS emerges, then, as a keystone facility because it is the only service which is open to the entire population of biological scientists and which provides bibliographic services covering most of the bio-science topics. It is not surprising then to find that requesters are asked to voluntarily restrict their demands (e. g., try other services first) or that service is relatively slow (up to 30 days to fill a request).

Two other government-sponsored facilities could offer services comparable to those of the NLM but are restrictive in coverage and access. The Defense Documentation Center is only moderately restrictive and could probably serve most of the bio-science community in some topic areas. However, while potential eligibility is broad, access requires a prior administrative procedure which could discourage some potential patrons. The National Agriculture Library appears substantially more restrictive with regard to both coverage and access and consequently does not provide a true alternative to the NLM-MEDLARS service.

With respect to access procedures, one fact stands out: the three facilities (CFSTI, DDC, and NLM-MEDLARS) which are on the low end of the scale of restrictiveness in eligibility, coverage, and scope of

service are the only three which require the submission of standard request forms; all other facilities are accessible to the eligible by informal, direct means.

Based on the results of this survey, it does not appear that formal relationships with existing information systems would increase the effectiveness or the variety of services that could be offered by the pilot information clearinghouse. It appears that continued exploitation of manual means for filling requests would be most efficient for the present.

2. Attitude of Biological Researchers

Only five (or 10%) of the scientists who were interviewed reported that a primary source of information was one or more of the special services available in the Washington, D. C. area. Most of those interviewed (60%) rely primarily on secondary sources such as abstracts, indices, copies of title pages of current journals, reviews, and incidental perusal of primary sources such as journals to which they subscribe or that are routed around the group of which they are a part. Another large segment (30%) do their own bibliographic work--relying on their own collections and the organized collections in the conventional library facilities provided by the organization within which they work.

Those relying predominantly on secondary sources report that "keeping abreast" or maintaining general currency regarding developments in their fields is an important objective. Documents such as Current Contents which give lists of titles of articles in a broad range of journals are particularly attractive to this segment. Seventeen of the fifty scientists interviewed indicated that this secondary publication was their most frequently used source of information.

Criticisms of the general characteristic of the "information environment" experienced by these bio-scientists were rare. Most often expressed dissatisfactions concerned details of coverage, delay, and cost, but not the basic attributes of information facilities and resources. Little interest was expressed in the processes of information service. For example, only one respondent remarked on the explicit use of inter-library loan procedures.

Based on the information obtained in this study, it may be concluded that very few basic biological researchers use existing information systems; some of them are not even aware of the existence of these systems.

One of the most important findings of this survey with respect to future clearinghouse studies is the emphasis placed on "keeping abreast" by the majority of these researchers. Although Current Contents, Chem. Abstracts, Biological Abstracts, and Index Medicus are all used for this

purpose, they are not entirely satisfactory. Perhaps a feature that could be evaluated in a clearinghouse study would be the impact of collecting such secondary source material based on a scientist's interest profile.

REFERENCES

Cuadra, C. A. (Ed.) Annual review of information science and technology. 2, New York: A. D. I., and John Wiley and Sons, 1967.

Herner, S. and Herner, M. Information needs and uses in science and technology. In Cuadra, C. A. op. cit.

Overhage, C. F. J. Plans for Project Intrex. Science, 1966, 152, 1032-1037.

Paisley, W. J. The flow of (behavioral) science information - A review of the literature. Institute for Communication Research. Palo Alto, California: Stanford University, November, 1965.

Recommendation for National Document Handling Systems in Science and Technology, Appendix A - A background study - vol. II Committee on Scientific and Technical Information (COSATI). Washington, D. C.: U. S. Dept. of Commerce, November, 1965.

Rubinoff, M. (Ed.) Toward a National Information System. Washington, D. C.: Spartan Books, Inc., 1965.

Salton, G. and Lesk, M. E. Computer evaluation of indexing and test processing. Information Storage and Retrieval. Scientific Report No. ISR-12. Department of Computer Science, Ithaca, New York: Cornell University, June 1967.

Schecter, G. (Ed.) Information Retrieval: A Critical View. Washington, D. C.: Thompson Book Company, 1967.

Swanson, D. R. On Improving Communication, Bulletin of the Atomic Scientists. February, 1966, 8-12.

Taylor, R. S., The Process of Asking Questions. American Documentation, 1962, October, 391-396.

U. S. Department of Health, Education and Welfare, Public Health Service. NINDB Neurological Information Network, Annual Progress Report, Bethesda, Maryland. October 1966.